



202173  
B

PATENT  
2080-3-32  
Customer No: 035884

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:

Sung Bae Jun et al.

Serial No: 09/932,713

Filed: August 17, 2001

For: METHOD AND APPARATUS FOR SKIMMING VIDEO DATA

Art Unit: 2173

Examiner: Nguyen, Cao H.

TRANSMITTAL OF PRIORITY DOCUMENT

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to:  
Assistant Commissioner for Patents  
Commissioner for Patents

P.O. Box 1450  
Alexandria, VA 22313-1450, on

July 13, 2005

Date of Deposit

Robert E. Kasody

Name

Signature

07/13/2005

Date

Dear Sir:

Enclosed herewith is a certified copy of Korean Patent Application No. 2000-48036 and 2000-55886 filed August 19, 2000 and September 22, 2000 and from which priority is claimed under 35 U.S.C. Section 119 and Rule 55.

Acknowledgment of the priority document(s) is respectfully requested to ensure that the subject information appears on the printed patent.

Dated: July 13, 2005

Respectfully submitted,

Customer No. 035884

By:

Robert E. Kasody  
Registration No. 50,268  
Attorney for Applicant(s)

801 S. Figueroa Street, 14th Floor  
Los Angeles, California 90017  
Telephone: (213) 623-2221  
Facsimile: (213) 623-2211



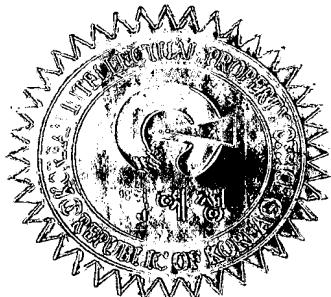
별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto is a true copy from the records of the Korean Intellectual Property Office.

출원 번호 : 특허출원 2000년 제 55886 호  
Application Number PATENT-2000-0055886

출원 년 월 일 : 2000년 09월 22일  
Date of Application SEP 22, 2000

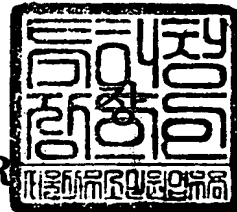
출원인 : 엘지전자주식회사  
Applicant(s) LG ELECTRONICS INC.



2001      년      10      월      23      일

특      허      청

COMMISSIONER



## 【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【제출일자】	2000.09.22
【국제특허분류】	H04N
【발명의 명칭】	비디오 콘텐츠의 구조적 정보를 이용한 비디오 스키밍 방법과 장치
【발명의 영문명칭】	A METHOD AND APPARATUS FOR VIDEO SKIMMING USING STRUCTURAL INFORMATION OF VIDEO CONTENTS
【출원인】	
【명칭】	엘지전자 주식회사
【출원인코드】	1-1998-000275-8
【대리인】	
【성명】	허용록
【대리인코드】	9-1998-000616-9
【포괄위임등록번호】	1999-043458-0
【발명자】	
【성명의 국문표기】	전성배
【성명의 영문표기】	JUN,Sung Bae
【주민등록번호】	711010-1057913
【우편번호】	153-034
【주소】	서울특별시 금천구 시흥4동 804
【국적】	KR
【발명자】	
【성명의 국문표기】	윤경로
【성명의 영문표기】	Y00N,Kyoung Ro
【주민등록번호】	641204-1037516
【우편번호】	135-271
【주소】	서울특별시 강남구 도곡1동 역삼 MBC 한신아파트 3동 1205호
【국적】	KR

**【발명자】****【성명의 국문표기】**

강배근

**【성명의 영문표기】**

KANG, Bae Guen

**【주민등록번호】**

600407-1916926

**【우편번호】**

139-200

**【주소】**서울특별시 노원구 상계동 172번지 대림아파트  
106-1308호**【국적】**

KR

**【심사청구】**

청구

**【취지】**특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규정에 의한 출원심사를 청구합니다. 대리인  
허용록 (인)**【수수료】****【기본출원료】**

20 면 29,000 원

**【가산출원료】**

14 면 14,000 원

**【우선권주장료】**

0 건 0 원

**【심사청구료】**

16 항 621,000 원

**【합계】**

664,000 원

**【첨부서류】**

1. 요약서·명세서(도면)\_1통

**【요약서】****【요약】**

본 발명은 멀티미디어 검색과 브라우징을 위한 시스템으로서, 특히 비디오 콘텐츠의 구조적 정보(샷 세그멘테이션 과정과 샷 클러스터링의 결과물인 비디오 콘텐츠의 구조적 정보)를 이용해서 비디오의 전체 내용에 대한 요약적인 이해와 사용자가 원하는 부분(위치)으로의 빠른 이동을 할 수 있도록 한 비디오 스키밍(skimming) 방법과 그 장치에 관한 것이다.

특히 본 발명은 샷 클러스터링(shot clustering)과 샷 세그멘테이션(shot segmentation)에 의해서 이루어지는 썸 및 샷을 기반으로 하여 비디오의 내용을 스키밍하고, 비디오 스키밍을 수행할 때 재생할 썸과 스킵할 썸을 선정한 다음, 재생할 썸내의 샷 내에서 특정한 일부분의 연속적인 재생 또는 스킵 기법을 도입한 부분 재생을 통해서 사용자가 짧은 시간 내에 비디오 전체의 내용을 충분히 이해할 수 있도록 함과 함께, 비디오 스트림을 빠른 시간내에 탐색하여 사용자가 원하는 위치로 빠르게 이동할 수 있도록 한 비디오 스키밍 방법과 그 장치에 관한 것이다.

**【대표도】**

도 10

**【색인어】**

비선형적 브라우징, 하이라이트, 스키밍, 비디오 인덱싱

【명세서】

【발명의 명칭】

비디오 콘텐츠의 구조적 정보를 이용한 비디오 스키밍 방법과 장치{A METHOD AND APPARATUS FOR VIDEO SKIMMING USING STRUCTURAL INFORMATION OF VIDEO CONTENTS}

【도면의 간단한 설명】

도1은 샷 세그멘테이션과 클러스터링의 개념을 설명하기 위한 도면

도2는 샷 세그멘테이션 정보를 이용한 비디오 스키밍 방법의 개념을 설명하기 위한 도면

도3은 대화 장면 썸의 샷 전개방식의 예를 나타낸 도면

도4는 샷 특성을 이용한 썸 검출방법의 예를 나타낸 도면

도5는 구조적 정보를 이용한 스키밍에서 재생할 샷과 스킵할 샷의 선정방법의 예를 나타낸 도면

도6은 썸 내부에서 샷의 위치와 반복정보를 고려한 스킵할 샷과 재생할 샷의 선정방법을 설명하기 위한 도면

도7은 샷 내에서 스킵할 부분과 재생할 부분의 선정방법을 설명하기 위한 도면

도8은 샷의 비유사도 특성을 이용한 동적인 단위 재생길이 선정방법의 예를 나타낸 도면

도9는 스키핑을 이용한 고속 스키밍 방법을 설명하기 위한 도면

도10은 비디오 콘텐츠의 구조적 정보를 이용한 스키밍 방법을 설명하기 위한 도면

도11은 비디오 콘텐츠의 구조적 정보를 이용한 비디오 스키밍을 위한 시스템 구성의 예를 나타낸 도면

**【발명의 상세한 설명】**

**【발명의 목적】**

**【발명이 속하는 기술분야 및 그 분야의 종래기술】**

<12> 본 발명은 멀티미디어 검색과 브라우징을 위한 시스템으로서, 특히 비디오 콘텐츠의 구조적 정보 중에서, 실제로 비디오의 내용 전개에 따른 의미있는 이야기 단위 구조인 씬을 기반으로 하여 비디오의 전체 내용에 대한 요약적인 이해와 사용자가 원하는 부분으로의 빠른 이동을 할 수 있도록 한 비디오 스키밍 시스템에 관한 것이다.

<13> 특히 본 발명은 멀티미디어 검색과 브라우징을 위한 시스템으로서, 샷 클러스터링의 결과물인 비디오 콘텐츠 구조적 정보로서 비디오 내용전개의 의미있는 이야기 단위인 씬 내에서 그 씬을 이루고 있는, 비디오 콘텐츠의 또다른 구조적 정보로서 물리적 단위인 샷들 중에서 재생할 샷과 스킵할 샷을 선정하고, 재생할 단위로 선정된 샷을 해당 씬의 내용 전개상의 실질적 의미있는 비디오로서 재생함으로써, 비디오의 전체 내용에 대한 요약적인 이해와 사용자가 원하는 부분(위치)으로의 빠른 이동을 할 수 있도록 한 비디오 스키밍(skimming) 방법과 그 장치에 관한 것이다.

<14> 더욱 상세하게는 본 발명은 샷 클러스터링(shot clustering)과 샷 세그멘테이션(shot segmentation) 과정에 의해서 이루어지는 썸 및 샷을 기반으로 하여 비디오의 내용을 스키밍하고, 비디오 스키밍을 수행할 때 재생할 썸과 스킵할 썸을 선정한 다음, 재생할 썸내의 샷 내에서 특정한 일부분의 연속적인 재생 또는 스킵 기법을 도입한 부분 재생을 통해서 사용자가 짧은 시간 내에 비디오 전체의 내용을 충분히 이해할 수 있도록 함과 함께, 비디오 스트림을 빠른 시간내에 탐색하여 사용자가 원하는 위치로 빠르게 이동할 수 있도록 한 비디오 스키밍 방법과 그 장치에 관한 것이다.

<15> 대중매체의 발달과 멀티미디어 콘텐츠의 제작이 쉬워짐에 따라 일반인이 매일 접하게 되는 미디어의 양이 매우 방대해 졌다. 멀티미디어 콘텐츠가 방대해짐에 따라 사용자가 원하는 데이터를 선별하여 주는 자동화 시스템에 대한 요구가 발생하였으며, 이를 해결하기 위한 방안들에 대한 연구가 활발히 진행되고 있다. 특히, 디지털 기술이 발달함에 따라 비디오 콘텐츠는 디지털 형태로 저장되어 유통되는 추세로 진행하고 있으며 디지털 방송이 대중화되면 이러한 미디어의 디지털화는 더욱 가속화 될 것이다.

<16> 이와같은 디지털 비디오 콘텐츠에 대하여 어떤 사용자는 뉴스에서 스포츠 관련 뉴스만을 시청하기를 원할 수 있으며, 또 다른 시청자는 뉴스에서 증권 관련 뉴스만을 원하는 경우가 있을 수 있다. 또한 어떤 사용자는 쇼 프로그램에서 특정 인물이 출연하는 장면만을 시청하기 위한 요구가 있을 수도 있다. 이러한 다양한 사용자의 요구를 수용하기 위하여 여러가지 연구가 활발히 진행되고 있다



<17> 또한 사용자는 제한된 시간 내에 비디오 콘텐츠의 전체 내용을 파악하고자 하는 요구가 있다. 이러한 요구를 수용하는 것이 '하이라이트'이다. 일반적으로 하이라이트는 비디오 콘텐츠에서 중요한 장면을 모아 새로 구성한 콘텐츠로 이해될 수 있으며, 이는 '스포츠 하이라이트', '영화 예고편', '헤드라인 뉴스' 등을 통하여 접할 수 있다. 그러나, 비디오 콘텐츠에서 하이라이트를 추출하는 것은 현재의 기술로 자동화 되기 매우 어려운 영역이다. 따라서 대부분 사람의 수동 작업에 의존하는 것이 보통이다. 앞서 언급된 바와 같이 미디어의 양이 폭발적으로 증가함에 따라 모든 비디오 콘텐츠에 대하여 하이라이트를 수동으로 제공하는 것은 많은 인력이 필요하게 되므로 거의 불가능하다. 따라서 사용자 하여금 빠른 시간내에 대강의 내용을 이해 시킬 수 있는 자동화 시스템이 필요하다.

— <18> 디지털 기술이 발달함에 따라 비디오 콘텐츠에서 사용자가 원하는 위치로 이동하고자 할 때 사용하기 위한 용도로 키프레임(key frame)이 이용된다. 키프레임을 이용한 비디오 요약에 이용하면 사용자는 원하는 곳으로 빨리 이동할 수 있는 장점이 있다. 그러나 키 프레임을 이용하여 사용자가 원하는 곳을 쉽게 찾기 위해서는 많은 수의 키 프레임이 필요하며, 제한된 디스플레이 공간에 많은 수의 키프레임을 디스플레이 하기 어려우므로 사용자에게 많은 선택 작업을 요구하게 되어 매우 불편하다. 또한 키프레임을 이용한 방법을 통해서도 일반적으로 전체 비디오의 내용을 이해하기 어려운 단점이 있다.

<19> 근래에는 디지털 비디오에서 원하는 장면을 찾기 위한 용도로 여러가지 비디오 인덱싱 기술이 연구되고 있다. 특정한 사람이 나온 장면만을 원하는 사용자를 위하여 비디오에서 사람이 나온 장면을 찾아내고 그 사람이 누구인지를 인식

하는 과정을 통하여 비디오에서 인물의 등장에 대한 정보를 인덱싱 하거나, 영화나 스포츠등에서 주요 장면을 추출하고 이를 인덱싱하는 등의 연구가 진행되고 있다. 그러나 비디오는 그 장르가 매우 다양하고 장르별로 인덱싱 되어야 할 데이터들이 매우 다르며, 현재의 기술로서는 사용자에게 의미 있는 정보를 높은 수준의 정확도를 가지고 추출할 수 있는 자동화 시스템의 구현은 매우 어려운 것으로 알려져 있다.

<20>        한편, 아날로그 비디오에서와는 달리 디지털 비디오에서는 빨리감기/빨리되감기 기능을 수행함에 있어서 화질의 저하를 막을 수 있다.

<21>        디지털 비디오에서 일반적으로 사용되는 고속 재생의 방법은 단위시간당 디코딩하는 프레임의 수를 늘려서 그 중의 일부분 만을 디스플레이 하거나 일정 부분을 건너뛰며 프레임을 디코딩하여 디스플레이 하는 방법이 사용된다.

<22>        그러나 단위시간당 디코딩하는 프레임의 수를 늘리는 방법은 단말 장치의 성능에 따라 최대 속도가 영향 받는 단점이 있으므로, 일반적으로 디지털 비디오의 빨리감기/빨리되감기에서는 일정 부분을 건너뛰며 프레임을 디코딩하여 디스플레이하는 방법을 이용한다. 디지털 비디오에서의 빨리감기/빨리 되감기 기술은 기존의 기술 중에서 제한된 시간내에 전체의 내용을 파악하고자 하거나 원하는 위치로 이동하고자 하는 사용자의 요구에 대응하기 위한 가장 합리적인 대안이지만, 일정부분을 스킵함에 있어서 대부분 시간적으로 일정한 간격을 이용하므로 사용자는 자신이 원하는 부분의 장면을 놓친다거나, 중요하지 않은 부분이 상대적으로 많이 재생되는 단점이 있다.

**【발명이 이루고자 하는 기술적 과제】**

- <23> 본 발명에서는 디지털 비디오 환경에서 사용자가 제한된 시간내에 비디오 전체의 내용을 파악하고자 하거나 또는 원하는 위치로 이동하고자 하는 사용자의 요구에 대응하며, 비디오에서 중요하지 않은 부분이 상대적으로 많이 재생된다거나 사용자가 실제로 시청을 원하는 장면을 놓칠 수 있는 확률을 최소화 하는 자동화된 비디오 스키밍 시스템 구성 방법을 제안한다.
- <24> 본 발명은 비디오 콘텐츠에 존재하는 논리적인 이야기 구조인 씬을 기반으로 하여 비디오 스키밍을 수행하되 씬을 이루고 있는 샷들 중에서 특정 샷을 선택하여 그 샷 또는 샷의 일부를 그 씬의 내용을 대표하는 정보로서 재생함으로써, 그 비디오의 실질적인 논리적 내용전개를 요약하여 일부로 전체 내용을 파악하거나, 또는 원하는 위치로 이동하고자 하는 사용자의 요구에 대응할 수 있도록 한 비디오 스키밍 시스템을 제안한다.
- <25> 특히 본 발명은 씬 정보와 샷 정보를 포함하는 비디오 콘텐츠의 구조적 정보를 이용해서 특정 이벤트가 교대로 반복하여 나타나는 장면에 대하여 보다 압축적인 정보를 제공함으로써, 보다 높은 수준의 비디오 검색과 브라우징을 가능하게 하는 비디오 스키밍 시스템을 제안한다.
- <26> 또한 본 발명은 알려진 바와같이 높은 정확도를 가지고 자동화 될 수 있는 비디오 세그멘테이션과 샷 클러스터링 기술에 기반하여 비디오 스키밍을 수행함을 특징으로 하며, 비디오의 씬 정보와 샷 정보를 이용해서 비디오 스키밍을 수행함으로써 사용자가 비디오의 전체 내용을 모두 시청하지 않고도 비디오의 내용을 이해하는데 필요충분할 정도로 요약된 검색 및 브라우징 환경을 제공함과 동

시에, 사용자가 원하는 위치로 빠르게 이동할 수 있는 비디오 스키밍 시스템을 제안한다.

**【발명의 구성 및 작용】**

<27> 본 발명의 비디오 콘텐츠의 구조적 정보를 이용한 비디오 스키밍 방법은, 비디오 스트림에 대한 구조적 정보로서 논리적인 이야기 구조의 단위인 씬(scene)에 대하여 씬의 구성요소인 물리적인 편집 단위인 샷(shot)들 중에서 재생할 샷과 스킵할 샷을 결정하는 샷 선정단계와, 상기 재생하기로 선정된 각 샷에 대하여 스킵할 부분과 재생할 부분을 결정하는 세그먼트 선정단계와, 상기 선정된 샷의 선정된 재생할 부분을 연속시켜 재생하는 단계; 를 포함하여 비디오 콘텐츠의 구조적 정보를 기반으로 하는 비디오 스키밍이 이루어지는 것을 특징으로 하는 비디오 스키밍 방법이다.

<28> 또한 본 발명에서 상기 비디오 스트림에 대한 구조적 정보는 논리적인 이야기 구조의 단위(story unit)인 씬(scene) 정보와 물리적인 편집단위인 샷(shot) 정보가 시간적인 기술 정보(시작위치와 지속시간 또는 시작위치와 종료위치)와 함께 표시된 것이고, 여기에 샷의 특성에 대한 기술 정보가 더 포함될 수 있음을 특징으로 하는 비디오 스키밍 방법이다.

<29> 또한 본 발명에서 상기 샷 선정 단계에서는, 샷의 특성이 비슷한 샷 중에서 일부분을 스킵하고 나머지 일부분의 샷들 만을 스키밍을 위하여 사용할 것으로 선정함으로써, 비슷한 특성을 가진 샷들이 반복적으로 재생되는 효과를 방지하는 것을 특징으로 하는 비디오 스키밍 방법이다.

- <30> 또한 본 발명에서 상기 유사한 샷들 중에서 재생에 이용될 샷을 선정함에 있어서, 썸의 후반부에 위치한 샷일수록 선택을 위한 높은 가중치를 부여하는 방식을 이용하여 스키밍에 이용할 샷을 선정하는 것을 특징으로 하는 비디오 스키밍 방법이다.
- <31> 또한 본 발명에서 상기 각각의 샷 내에서 재생할 부분의 선정은 해당 샷의 전반부, 후반부, 중간부분, 전반부와 후반부를 동시에 사용하는 것이 가능함을 특징으로 한다.
- <32> 또한 본 발명에서 상기 각각의 샷에서 재생될 길이의 결정은 각각의 샷에서 동일한 길이의 세그먼트를 선택하여 재생하거나, 개별 샷내의 이미지적/모션적/오디오적 유사성의 평균값에 근거하여 유사도가 높으면 재생할 길이를 줄이고 유사도가 낮으면 해당 샷에서 재생할 길이를 늘리는 것을 특징으로 하는 비디오 스키밍 방법이다.
- <33> 또한 본 발명에서 상기 샷 내의 이미지적/모션적/오디오적 유사성은 샷 내에서의 시간적 위치가 다른 프레임, 움직임 벡터, 오디오 데이터의 유사성인 것을 특징으로 하는 비디오 스키밍 방법이다.
- <34> 또한 본 발명에서 상기 개별 샷에서 재생할 부분으로 선정된 세그먼트의 길이가 해당 샷의 길이보다 크게 계산될 경우 개별 샷에서 재생할 부분의 길이를 해당 샷의 길이 이하로 줄이는 것을 특징으로 하는 비디오 스키밍 방법이다.
- <35> 또한 본 발명에서 상기 비디오 스키밍을 위해 선택되어 재생될 구간들을 정상적인 비디오 재생속도로 재생하거나, 또는 단위 시간당 디코딩될 프레임의 수

를 정상적인 경우보다 높여서 상기 재생구간을 고속 재생하거나, 또는 상기 재생할 구간내의 프레임들을 모두 디코딩하지 않고 중간의 몇 프레임씩 스킵하여 재생함으로써 상기 재생구간을 고속 재생하는 것을 특징으로 하는 비디오 스키밍 방법이다.

<36> 또한 본 발명에서 상기 스킵을 이용한 고속 스키밍을 MPEG과 같은 프레임간 압축을 사용하는 코딩 스킴을 이용한 비디오 스트림에 적용하였을 때, 디코딩할 프레임은 다른 프레임의 디코딩 없이 해당 프레임만 디코딩하여 프레임 데이터를 얻을 수 있는 I 프레임인 것을 특징으로 하는 비디오 스키밍 방법이다.

<37> 한편, 본 발명의 비디오 콘텐츠에 대한 구조적 정보를 이용한 비디오 스키밍 장치는, 멀티미디어 데이터로서 디지털 비디오 데이터의 검색과 브라우징이 이루어지기 위하여 비디오 스키밍을 위한 사용자 명령을 입력하는 사용자 인터페이스 수단과, 상기 사용자 인터페이스수단으로 입력된 사용자 명령에 따라 해당 비디오 화일을 비디오 콘텐츠에 대한 구조적 정보로서 썸정보 및 샷 정보에 근거하여 스키밍 처리하는 제어수단과, 상기 제어수단에 디지털 비디오 데이터와 해당 비디오에 대한 인덱스 정보로서 비디오 콘텐츠에 대한 썸정보 및 샷정보를 제공하기 위한 비디오 정보 화일과, 상기 제어수단에 의해서 스키밍 처리된 비디오가 재생되는 디스플레이수단을 포함하여 이루어진 것을 특징으로 하는 비디오 스키밍 장치이다.

<38> 또한 본 발명의 비디오 스키밍 장치에서, 상기 사용자 인터페이스 수단은 비디오 스키밍의 정도로서 요약 수준을 지정하는 수단이나, 비디오 스키밍시의 재생 구간의 배속을 지정하는 수단을 포함하여, 비디오 스키밍이 이루어질 때 비

디오의 요약수준이나 재생 배속을 선택할 수 있는 것을 특징으로 하는 비디오 스키밍 장치이다.

<39> 또한 본 발명의 비디오 스키밍 장치에서, 상기 제어수단은 사용자의 입력 또는 기본 설정을 이용하여 스키밍 조건에 따라 인덱스 화일로부터 샷세그멘테이션 정보 및 샷 클러스터링 정보와 관련된 비디오 인덱스정보를 읽고, 비디오 스키밍 조건에 맞게 재생할 세그먼트들을 계산하여 관련된 미디어 화일에서 해당 세그먼트들을 연속적으로 재생하여 디스플레이 수단으로 출력하는 것을 특징으로 하는 비디오 스키밍 장치이다.

<40> 상기한 바와같이 이루어진 본 발명의 비디오 콘텐츠의 구조적 정보인 썸과 샷 정보를 이용한 비디오 스키밍 시스템을 첨부된 도면을 참조하여 더욱 상세하게 설명한다.

<41> 디지털 비디오 기술의 발달과 이미지/비디오 인식 기술의 발달로 인하여 사용자들은 원하는 비디오를 원하는 시점에 원하는 부분만을 검색(searching/filtering)하고 브라우징(browsing) 할 수 있게 되었다.

<42> 비선형적인 비디오 브라우징(non-linear video browsing)과 검색을 위하여 가장 기본이 되는 기술은 샷 세그멘테이션 (shot segmentation) 기법과 샷 클러스터링(shot clustering) 기법이며, 이 두 가지 기술은 비디오를 분석(analysis)하는데 있어서 가장 핵심이 되는 기술이다. 따라서 현재까지 많은 연구가 샷 세그멘테이션에 집중되어 있고 샷 클러스터링 기술에 대한 연구가 시작되고 있는 추세이다.

- <43> 여러가지 연구 결과를 토대로 하면 샷 세그멘테이션은 자동화 될 수 있으며 대부분의 알고리즘은 90%이상의 높은 정확도를 가지고 구현될 수 있다.
- <44> 또한 샷 클러스터링 기술도 특징적인 이벤트를 검출하거나 일반적인 샷의 특성을 이용하여 프로그램의 장르에 맞도록 기술을 적용함으로써 높은 수준의 정확도를 가지고 자동화될 수 있다.
- <45> 일반적으로 비디오 콘텐츠는 논리적으로 몇 개의 이야기 구조의 단위(Story Unit)로 나뉘어 진다. 이러한 이야기 구조의 단위를 일반적으로 사건구조(event) 또는 씬(scene)이라고 일컫는다. 총격전 장면이라든지, 대화 장면 등이 이에 해당한다. 이러한 씬들은 여러 개의 서브 씬 또는 샷들의 연결로 구성된다.
- <46> 샷(shot)이란 방해(interruption) 없이 하나의 카메라로부터 얻어진 비디오 프레임들의 시퀀스(sequence)를 의미하며, 이는 비디오를 분석(analysis)하거나 구성(construction)하는 가장 기본이 되는 단위이다.
- <47> 일반적으로 비디오는 수많은 샷의 연결로 구성된다. 샷 세그멘테이션이란 비디오를 각각의 개별 샷으로 분할하는 기법을 의미하며, 샷 클러스터링(shot clustering)은 비디오 콘텐츠를 각각의 개별 샷과 그 특성에 기반하여 샷들을 논리적인 씬 단위로 재구성함으로써 비디오 콘텐츠의 논리적인 이야기 구조를 검출하는 과정을 의미한다.
- <48> 도1은 샷 세그멘테이션 과정과 샷 클러스터링 과정을 도식화 한 것이다. 일반적으로 대부분의 샷 세그멘테이션 알고리즘은 동일한 샷내에서는 이미지적/모션적/오디오적 유사성이 존재하고 두개의 다른 샷 간에는 이미지적/모션적/오디



오적 비유사성이 계측되는 특성에 기반하며, 샷 클러스터링은 유사한 특성을 가진 샷들이 일정 시간 내에 다시 검출되는 특징에 기반한다.

<49> 일반적으로 비디오 하이라이트는 비디오 스트림의 내용전개에 있어서 의미상으로 중요한 세그먼트들을 선택하여 이 세그먼트들을 연속적으로 재생하는 방법이다.

<50> 그러나 다양한 비디오 콘텐츠에 대하여 내용전개에 있어서 의미상으로 중요한 세그먼트들을 선택하는 것은 자동화 하기 매우 어렵다.

<51> 그런데, 샷 세그멘테이션 정보를 비디오 스키밍에 이용한다면, 모든 비디오에 존재하는 샷을 기준으로 하여 샷의 일정부분만을 재생하고 나머지 부분은 스킵하는 방법을 통하여 원 스트림 보다 짧은 길이를 재생하도록 하는 스키밍 방법을 구사할 수 있다. 이러한 스키밍 방법은 샷 세그멘테이션 기술이 자동화 가능하므로 완전 자동화 시스템을 구축할 수 있다는 장점이 있으며, 일반적인 디지털 비디오에 대한 빨리감기/빨리 되감기에서 발생하는 중요하지 않은 장면이 길게 재생된다거나 중요한 장면을 놓친다거나 하는 문제를 줄일 수 있는 장점이 있다.

<52> 도2는 샷 세그멘테이션 정보를 이용한 비디오 스키밍 방법을 요약한 도면이다. 도2에서 회색으로 표시된 부분은 샷 세그멘테이션 정보를 이용한 스키밍 방법에서 재생될 부분을 나타낸 것이며, 나머지 부분은 스킵할 부분을 나타낸 것이다.

- <53> 그러나 샷 세그멘테이션 정보만을 비디오 스키밍에 이용할 경우에는 비디오 콘텐츠에 존재하는 논리적인 이야기 구조인 씬 정보를 이용하지 않기 때문에, 대화장면 등과 같은 특정 사건구간에서는 반복적인 샷들이 계속하여 플레이 되는 단점이 있다.
- <54> 도3은 대화장면이 길게 연결되는 씬에서의 샷의 배열 구조를 도식화 한 것이다. 도3에서 각각의 샷은 샷 세그멘테이션 과정을 통하여 검출된 샷 특성에 기반하여 영어 대문자(A,B,C,D)로 표현하였다.
- <55> 즉, 도3에서 표현한 대화 장면은 등장인물1과 등장인물2가 교대로 클로уз업(Closed-up)되어 이야기를 전개해 나가는 씬으로서 많은 샷들로 구성되어 있다.
- <56> 그러나 샷 세그멘테이션 정보만을 비디오 스키밍에 이용해서는 대화장면 내부의 각 샷들의 일부분을 모두 재생 하므로, 사용자에게는 두사람이 대화를 한다는 정보 이외에 크게 다른 정보를 추가적으로 제공할 수 없음에도 불구하고 이 씬이 오랫동안 재생되는 단점이 있다.
- <57> 본 발명에서는 비디오 콘텐츠의 구조적 정보로서 씬 정보와 함께 샷 정보를 고려하여 비디오 스키밍을 수행함으로써, 위와같은 단점을 극복한다.
- <58> 즉, 본 발명에서는 모든 비디오에 존재하는 씬에서 그 씬을 이루고 있는 샷들 중에서 재생할 샷과 스킵할 샷을 먼저 가려내고, 재생할 샷에 대해서는 그 샷을 이루고 있는 세그먼트 정보 중에서 샷의 일정부분(세그먼트)만을 재생하고 나

머지 부분은 스킵(skip)하는 방법을 통하여 원래의 비디오 스트림 보다 짧은 길이를 재생하도록 하는 스키핑 방법과 그 장치를 제안한다.

<59> 여러가지 연구 결과는 비디오 콘텐츠에서 충격장면, 대화 장면 등의 특정한 이벤트가 검출될 수 있음을 기반으로 하여 영화나 드라마와 같은 콘텐츠의 씬 검출이 가능하며, 이를 통하여 목차(ToC : Table of Content) 형태의 인덱스 구조를 자동으로 생성할 수 있음이 알려졌다.

<60> 도4는 일반적인 비디오 콘텐츠에 대한 이야기 구조 단위(Story Unit)를 검출하는 과정을 도식화 한 것이다.

<61> 도3에서와 마찬가지로 각각의 샷은 샷 세그멘테이션 과정을 통하여 검출된 샷 특성에 기반하여 영어 대문자로 표현하였다. 드라마나 영화의 대화장면이 있는 씬 내부의 샷 전개 구조에서는 대부분 A,B,A,B,...와 같은 특징적인 패턴이 나타난다. 도4에서는 일정한 주기 이내에 특성이 비슷한 샷이 검출되면 해당 구간을 하나의 씬으로 결정하는 과정을 보여 준다. 도4에서 씬1은 샷의 특징 값이 A,B,C인 샷들로 이루어진 씬이다. 샷1-B<sub>3</sub> 이후에 일정 시간동안 샷의 특징 값이 A,B,C인 샷이 존재하지 않으므로 샷1-B<sub>3</sub>의 종료시점을 씬1의 종료시점으로 검출하여 씬을 검출한다. 도4에서 씬2는 샷의 특징 값이 F,H,E인 샷들로 이루어진 씬이고, 이 씬의 마지막 샷 이후에 일정 시간동안 샷의 특징 값인 F,H,E가 존재하지 않으므로 씬2의 종료시점을 검출할 수 있다.

<62> 이러한 방법 이외에도 사람의 얼굴 부분을 자동으로 검출(Face Detection)하고 검출된 사람이 누구인지를 얼굴 인식(Face Recognition)하는 과정을 통하여

보다 정확한 대화 장면 검출이 가능하다. 이러한 방식은 일반적인 드라마나 영화에 잘 적용되는 방식이다.

<63> 본 발명은 앞에서 설명한 바와같이 비디오 콘텐츠에 대한 구조적 정보인 씬과 샷 정보를 이용해서 비디오 스키밍을 구현하며, 씬을 이루고 있는 개별 샷들 중에서 재생할 샷을 어떻게 선정할 것인가, 재생할 샷으로 선정된 샷 내에서 재생될 부분과 스킵할 부분을 어떻게 선정할 것인가, 재생할 부분의 재생길이는 어떻게 선정할 것인가, 그리고 재생구간 내에서의 재생방법은 어떻게 할 것인가를 고려한다.

<64> 먼저, 도5는 본 발명의 비디오 스키밍 방법을 요약한 도면이다.

<65> 도5에서는 샷 세그멘테이션과 샷 클러스터링 과정을 통하여 인덱스된 비디오 콘텐츠의 구조적 정보를 이용한다. 도5에서는 구조적 정보를 이용한 비디오 스키밍시에 재생을 위하여 선택된 씬들을 회색으로 표시하고 스킵될 씬들을 흰색으로 표시하였다. 즉, 구조적 정보를 이용한 비디오 스키밍을 위하여 시스템은 우선 각 씬에 대하여 재생될 샷들을 결정하고 개별 샷에 대한 재생 방법을 결정한다.

<66> 도5에서는 씬1의 각 샷들 중에서 반복적인 샷들을 재생하지 않기 위하여 비슷한 샷들 중에서 특징적인 장면만을 한번씩만 재생하도록 한 예제이다.

<67> 본 발명에서 비디오 스트림에 존재하는 각각의 씬(scene)에 대하여 씬의 구성요소인 샷(shot)들 중에서 재생할 샷과 스킵할 샷을 결정하는 샷 선정은 다음과 같이 이루어진다.

<68> 하나의 씬에 비슷한 특성의 샷이 여러 개 존재하는 경우에 대표 샷을 선정하는 방법은 특별한 가중치 조건 없이 선정하여 스키밍에서 사용하더라도 대강의 씬의 내용을 전달할 수 있다. 그러나 일반적인 드라마나 영화와 같은 이야기 구조에서는 한 씬의 후반부에 더욱더 많은 정보가 표현되는 경우가 대부분이다. 다시 말하면 도입 부분은 결말 부분보다 그 중요성이 일반적으로 떨어진다. 따라서 비슷한 샷들이 씬내에 여러 번 등장할 때 스키밍에서 재생할 샷들을 선정하는 단계에서는 샷의 후반부에 있는 샷들을 재생할 샷으로 선정하는 것이 같은 길이의 스키밍 방법에서 사용자에게 더욱더 많은 정보를 제공할 수 있는 방법이 된다.

<69> 도6은 샷의 전반부에서 스키밍에서 재생할 샷들을 선정하는 방법(도6의 a)과 샷의 후반부에서 재생할 샷들을 선정하는 방법(도6의 b)을 나타낸 것이다.

<70> 도6의 a,b는 모두 한 씬에서 비슷한 샷들이 존재하는 경우 하나의 샷만을 스키밍을 위하여 선정하는 예이며, 따라서 도6의 a에서는 A,B,C 각각의 샷 특성을 갖는 샷들 중에서 제일 처음 나오는 샷들을 재생할 샷으로 선정하였고, 도6의 b에서는 A,B,C 각각의 샷 특성을 갖는 샷들 중에서 제일 마지막으로 나오는 샷들을 재생할 샷으로 선정하였다. 일반적으로는 도6의 b의 방법이 도6의 a의 방법보다 더욱더 사용자의 만족도가 높게 나타난다.

<71> 다음, 각 샷 내에서 재생될 부분과 스킵핑할 부분을 선정하는 방법에 대해서 설명한다.

<72> 비디오 콘텐츠의 구조적 정보를 이용한 스키밍에서는 앞서 선정된 샷들을 연속적으로 재생함으로써 비디오 콘텐츠에 대한 요약を提供할 수 있다. 그러나 일반적으로 샷 전체를 플레이하는 비디오 스키밍 방식은 요약의 수준이 매우 낮

으며 사용자는 일반적으로 샷의 일부만을 시청하더라도 전체의 내용을 이해할 수 있다. 비디오 콘텐츠의 구조적 정보를 이용한 비디오 스키밍에서 재생을 위하여 선택된 샷에서 재생될 부분을 선정하는 방법은 무조건적으로 샷의 앞부분을 선택하거나 뒷부분 또는 가운데 부분을 선택할 수 있다. 도7은 샷의 앞부분, 뒷부분, 가운데 부분 또는 앞/뒤 부분을 동시에 이용하여 비디오 스키밍을 하는 방식에서 재생될 부분과 스킵될 부분을 나타낸 도면이다.

<73> 그러나 비디오의 장르에 따라서 차이는 있지만 실험의 결과는 해당 샷의 앞부분을 스킵하고 뒷부분을 재생하는 것이 일반적으로 사용자의 만족도가 높은 것으로 나타난다. 그 이유는 일반적으로 샷의 결말 부분(예 : 축구의 골 장면 등)이 샷의 내용을 이용하는데 있어서 도입이나 전개 부분보다 더 중요하며, 뉴스와 같은 프로그램에서 단계적인 도표 설명등의 방법을 이용할 경우 샷의 앞부분에서는 내용의 일부가 표현되고 마지막 부분에서 전체의 내용이 표현되기 때문이다.

<74> 그러나 비디오의 장르에 따라서는 샷의 앞부분이 일반적으로 중요한 것도 있는데 그 일례로 문제 풀이 등을 주로 하는 교육 방송등을 들 수 있다. 이러한 방송에서는 샷의 앞부분에서 어떤 문제를 다루는지에 대한 정보가 있고 그 이후로는 문제 풀이가 이어지므로 원하는 부분을 재생하기 위해서는 샷의 앞부분을 재생하는 것이 뒷부분을 재생하는 것보다 사용자에게 더 많은 정보를 제공할 수 있다.

<75> 따라서 본 발명에서는 비디오의 콘텐츠의 성격에 따라서 샷 내에서 재생할 위치를 다르게 선정할 수 있으며, 같은 샷 내에서도 앞부분, 중간부분, 뒷부분을 혼용하여 스키밍을 구현할 수 있다.

- <76>        다음은 본 발명에서 재생될 길이를 선정하는 방법에 대해서 설명한다..
- <77>        각 샷에서 재생될 길이를 선정하는 방법은 선정된 모든 샷에 대하여 동일한 길이의 세그먼트를 재생할 부분으로 선정하는 방법과, 샷의 특성을 이용하여 각 샷마다 재생할 길이를 다르게 선정하는 방법으로 나눌 수 있다.
- <78>        이때 이용하는 샷의 특성은 한 샷 내에서의 평균적 이미지적/모션적/오디오적 유사성에 근거할 수 있다. 즉 한 샷내에서 이미지적/모션적/오디오적 유사성이 크면 클수록 장면이 단조로운 것으로 판단할 수 있으므로 이러한 장면에 대해서는 스킵을 많이 하고, 샷내에서 이미지적/모션적/오디오적 유사성이 작으면 내용이 복잡한 장면으로 판단할 수 있으므로 이러한 장면에 대해서는 스킵을 적게 하는 방식을 이용하여 재생될 단위 세그먼트의 길이를 동적으로 조정할 수 있다.
- <79>        이러한 방법은 샷의 시간적 길이에 의존하지 않고 내용이 많은 부분은 적게 스킵하고 내용이 적은 부분은 많이 스킵하는 방법으로, 선정된 모든 샷에 대하여 동일한 길이의 세그먼트를 재생하는 방법보다 사용자의 이해도가 높은 수준의 비디오 스키밍을 제공할 수 있는 방법이다.
- <80>        도8은 샷내에서 이미지적/모션적/오디오적 유사성에 기반한 재생 및 스킵할 길이의 선정 방법의 일례를 나타낸 것이다.
- <81>        도8의 그래프에서 가로축은 시간을 의미하며 세로축은 샷 내에서의 이미지적/모션적/오디오적 비유사성을 측정하여 누적시킨 값을 의미한다. 이러한 비유

사성 데이터는 일반적으로 샷 세그멘테이션 알고리즘에서 추출 가능한 샷 특성을 나타내는 데이터이다.

<82> 비유사도의 일례로 인접한 프레임 또는 일정간격으로 떨어진 프레임간의 칼라 히스토그램(color histogram)의 차이를 들 수 있다.

<83> 도8에서는 샷A와 샷B는 길이가 비슷하지만 샷A의 평균 변화율 보다 샷B의 평균 변화율이 더 크므로 샷A에서보다 샷B에서 더 많은 부분을 재생하는 상황을 표현하였다.

<84> 이와같이 재생 구간을 설정함에 있어서 샷의 길이를 고려하지 않으면 해당 샷의 길이보다 재생할 구간의 길이가 더 커지는 예러 상황(샷의 길이가 매우 짧은 경우)이 발생할 수 있으므로 본 발명의 스키밍 방법에서는 단위 구간의 길이가 해당 샷 보다 커지는 경우에는 예외적으로 해당 샷의 전체를 재생 구간으로 선정하거나 해당 샷의 길이를 고려하여 일부분을 재생 구간으로 선정할 수 있다.

<85> 다음은 비디오 콘텐츠에 대한 구조적 정보로서 상기한 바와같이 썸과 그 썸 내의 재생할 샷 내에서의 재생 구간내에서의 재생 방법에 대하여 설명한다.

<86> 본 발명에 따른 비디오 스키밍은 순방향 뿐만아니라 역방향에 대하여서도 적용이 가능하다.

<87> 각각의 샷내에서 재생할 구간으로 선정된 세그먼트들을 연속적으로 재생하면 사용자는 전체의 내용을 이해하면서도 짧은 시간에 콘텐츠에 대한 개요정보를 얻을 수 있으며 원하는 위치를 탐색하는데 있어서 별도의 간섭이 필요하지 않다



- <88> 본 발명의 비디오 스키밍 방법에서 각각의 샷내에서 재생할 구간으로 선정된 세그먼트들을 재생하는 방법은 크게 두가지로 나뉘 수 있다.
- <89> 첫번째는 정상 재생 방법과 동일한 방법으로 각 세그먼트들을 재생하는 방법이며, 두번째는 재생 구간내의 일부 프레임만을 디코딩하여 구간내에서 다시 스킵핑을 이용하여 재생하는 방법이다.
- <90> 정상 재생 방법은 매우 일반적이므로 구체적인 설명을 생략하고 재생 구간내에서 일부 프레임만을 디코딩하여 구간내에서 스킵핑을 이용한 재생 방법을 설명하기로 한다.
- <91> 재생 구간내에서 일부 프레임만을 디코딩하여 구간내에서 스킵핑을 이용한 재생 방법은 고속 스키밍을 구현하기 위한 방법이다. 이때 디스플레이될 프레임들은 시간적으로 일정한 간격 만큼 떨어진 프레임으로 지정할 수도 있으며 MPEG과 같은 프레임간 압축을 이용하는 방법에서는 프레임간 의존성이 없는 I 프레임들로 지정할 수 있다.
- <92> 도9는 재생 구간내에서 스킵핑을 이용한 고속 스키밍 방법의 일례를 도식화한 것이다. 이러한 방법을 사용하면 사용자는 실제로 많은 정보를 얻으면서 고속으로 비디오 화일을 재생하는 효과를 경험할 수 있다.
- <93> 이상에서 설명한 바와같이 본 발명의 비디오 콘텐츠의 구조적 정보를 이용한 비디오 스키밍 방법은 두단계로 세그먼트를 지정하는 것이 그 특징이며, 도10은 이와같은 본발명 비디오 콘텐츠의 구조적 정보를 이용한 비디오 스키밍 방법을 요약한 도면이다.

<94> 비디오 스키밍이 요구되면 시스템은 비디오 콘텐츠에 대한 샷 및 씬정보를 포함한 비디오 콘텐츠의 구조적 정보가 저장된 인덱스 파일이 로드된다. 시스템은 각 씬에 대하여 어떤 샷을 재생할 것이고 어떤 샷을 스킵할지를 결정하고(샷 선정 단계), 비디오 스키밍을 위하여 선정된 각 샷에 대하여 재생할 세그먼트와 스킵할 세그먼트를 결정하고(세그먼트 지정 단계) 두 단계의 결정 과정을 통하여 재생할 세그먼트들을 연속적으로 재생 장치로 출력하게 된다.

<95> 도10에서는 첫번째의 샷 선정 단계를 거쳐서 재생할 샷들을 회색으로 표시하였으며, 선정된 샷에서 일부분만(세그먼트)을 재생하고 나머지 부분을 스킵할 수 있음을 도식화 하였다.

<96> 도11은 비디오 스키밍을 위한 본 발명의 스키밍 장치의 실시예를 나타낸 것이다.

<97> 도11에 의하는 바와같이 본 발명의 비디오 스키밍 장치는, 비디오 스키밍의 정도와 스키밍에서 사용할 배속 등의 사용자 명령을 입력하기 위한 사용자 인터페이스부(101)와, 상기 사용자 인터페이스부(101)로 입력된 사용자 명령에 따라 해당 비디오 화일을 샷 및 씬에 대한 인덱싱 정보에 근거하여 스키밍 처리하는 주제어부(102)와, 상기 주제어부(102)에 디지털 비디오 스트림 정보를 제공하기 위한 미디어 화일(103)과, 상기 미디어 화일에 해당하는 구조적 정보로서 씬 및 샷에 대한 인덱싱 정보를 제공하기 위한 인덱스 화일(104)과, 상기 주제어부(102)에 의해서 스키밍 처리된 비디오가 재생되는 디스플레이 장치부(105)를 포함하여 이루어진다.

<98> 도11의 본 발명 비디오 스키밍 시스템에서 인덱스 화일(104)은 미디어 화일(103)에 포함될 수 있다. 디스플레이 장치부(105)는 비디오 스트림을 디스플레이하는 모니터, 스피커등의 출력 장치이며, 사용자 인터페이스부(101)는 사용자의 입력을 받아들이는 키보드, 마우스, 리모콘, 버튼 등의 입력 수단을 의미한다.

<99> 미디어 화일(103)은 비디오(오디오)데이터가 저장된 화일이며, 인덱스 화일(104)은 샷 클러스터링 및 샷 세그멘테이션 정보가 포함된 비디오에 대한 인덱스 정보가 저장된 화일이다.

<100> 사용자는 사용자 인터페이스부(101)를 이용하여 비디오 스키밍을 요구한다.

<101> 비디오 스키밍을 요구할 때에는 요약의 수준(스키밍의 정도)를 지정할 수 있으며, 스키밍에서 사용할 배속을 지정할 수 있다. 즉 사용자는 전체 비디오를 몇 분안에 압축하여 시청할지를 사용자 인터페이스부(101)를 이용하여 지정하고, 주 제어부(102)는 사용자의 입력에 대응하여 미디어 화일(103)과 그에 따른 인덱스 화일(102)의 정보로부터 어떤 샷의 어떤 부분을 스키밍을 위하여 재생할 것인지, 각 세그먼트들은 몇 배속으로 재생할 것인지를 결정하게 된다. 이러한 과정을 마치면 주 제어부(102)는 미디어 화일(103)을 디코딩하여 해당 프레임들을 디스플레이 장치부(105)에 디스플레이함으로써 사용자에게 비디오 스키밍 기능을 제공하게 된다.

<102> 상기한 바와같이 본 발명은 디지털 비디오 환경에서 제한된 시간내에 전체의 내용을 파악하고자 하거나 원하는 위치로 이동하고자 하는 사용자의 요구에 동시에 대응하는 비디오 스키밍 방법을 제안하였다.

<103>      본 발명은 기존의 비디오 스키밍에서 발생할수 있는 문제점인, 중요하지 않은 부분이 상대적으로 많이 재생된다거나 사용자가 실제 원하는 장면을 놓칠 수 있는 확률을 최소화 하며, 대화장면이나 특정 장면이 교대로 반복되는 장면이 계속해서 재생될 확률을 최소화하였다.

<104>      본 발명의 비디오 스키밍 방식은 원하는 위치로 이동하고자 하는 사용자의 요구에 대하여 사용자 입력의 필요성을 최소화 시킬수 있는 방법이다.

**【발명의 효과】**

<105>      본 발명의 비디오 스키밍을 기능을 이용하면 사용자는 짧은 시간내에 전체의 내용을 파악할 수 있으며 전체의 내용을 파악함에 있어서 중요한 부분을 놓치지 않고 지루한 부분을 간단히 지나갈 수 있는 장점이 있다.

<106>      또한 사용자는 원하는 위치로 이동하고자 할 때에도 본 발명의 비디오 스키밍을 이용할 수 있으며 이는 키프레임을 이용한 방식보다 사용자 입력 요구가 매우 작은 장점이 있다.

<107>      결국 본 발명은 비디오 하이라이트 재생과 같은 용도로 이용될 수 있으며, 각 샷의 재생할 구간들을 재생함에 있어서 고속재생방식과 혼용되면 사용자의 입력 요구를 최소화 하며 원하는 장면을 빨리 탐색하는 기능으로 활용될 수 있다.

1020000055886

출력 일자: 2001/10/24

**【특허청구범위】****【청구항 1】**

비디오 스트림에 대한 구조적 정보로서 논리적인 이야기 구조의 단위인 씬(scene)에 대하여 씬의 구성요소인 물리적인 편집 단위인 샷(shot)들 중에서 재생할 샷과 스킵할 샷을 결정하는 샷 선정단계와, 상기 재생하기로 선정된 각 샷에 대하여 스킵할 부분과 재생할 부분을 결정하는 세그먼트 선정단계와, 상기 선정된 샷의 선정된 재생할 부분을 연속시켜 재생하는 단계;를 포함하여 씬을 기반으로 하는 비디오 스키밍이 이루어지는 것을 특징으로 하는 비디오 스키밍 방법.

**【청구항 2】**

제 1 항에 있어서, 상기 비디오 스트림에 대한 구조적 정보는 논리적인 이야기 구조의 단위(story unit)인 씬(scene) 정보와 물리적인 편집단위인 샷(shot)정보가 시간적인 기술 정보(시작위치와 지속시간 또는 시작위치와 종료위치)와 함께 표시된 것이고, 여기에 샷의 특성에 대한 기술 정보가 더 포함될 수 있음을 특징으로 하는 비디오 스키밍 방법.

**【청구항 3】**

제 1 항에 있어서, 상기 샷 선정 단계에서는, 특성이 비슷한 샷 중에서 일부분을 스킵하고 나머지 일부분의 샷들 만을 스키밍을 위하여 사용할 것으로 선정함으로써, 비슷한 특성을 가진 샷들이 반복적으로 재생되는 효과를 최소화하는 것을 특징으로 하는 비디오 스키밍 방법.

**【청구항 4】**

제 3 항에 있어서, 상기 유사한 샷들 중에서 재생에 이용될 샷을 선정함에 있어서, 썸의 후반부에 위치한 샷일수록 선택을 위한 높은 가중치를 부여하는 방식으로 스키밍에 이용할 샷을 선정하는 것을 특징으로 하는 비디오 스키밍 방법.

**【청구항 5】**

제 1 항에 있어서, 상기 썸을 대표하는 재생할 샷내의 그 샷을 대표하는 재생할 부분(세그먼트)은, 해당 샷의 전반부, 후반부, 중간부분, 전반부와 후반부를 동시에 사용하는 것을 특징으로 하는 비디오 스키밍 방법.

**【청구항 6】**

제 1 항에 있어서, 상기 썸을 대표하는 재생할 샷내의 그 샷을 대표하는 재생할 부분(세그먼트)의 길이는 각 샷에서 동일한 길이로 선택하여 재생하는 것을 특징으로 하는 비디오 스키밍 방법.

**【청구항 7】**

제 1 항에 있어서, 상기 썸을 대표하는 재생할 샷내의 그 샷을 대표하는 재생할 부분(세그먼트)의 길이는 개별 샷내의 이미지적/모션적/오디오적 유사성의 평균값에 근거하여 유사도가 높으면 재생할 길이를 줄이고 유사도가 낮으면 해당 샷에서 재생할 길이를 늘리는 것을 특징으로 하는 비디오 스키밍 방법.

**【청구항 8】**

제 7 항에 있어서, 상기 썸을 대표하는 샷내의 이미지적/모션적/오디오적 유사성은 샷 내에서의 시간적 위치가 다른 프레임, 움직임 벡터, 오디오 데이터의 유사성인 것을 특징으로 하는 비디오 스키밍 방법.

**【청구항 9】**

제 6 항 또는 제 7 항에 있어서, 상기 썸을 대표하는 개별 샷에서 재생할 부분으로 선정된 세그먼트의 길이가 해당 샷의 길이보다 크게 계산될 경우 개별 샷에서 재생할 부분의 길이를 해당 샷의 길이 이하로 줄이는 것을 특징으로 하는 비디오 스키밍 방법.

**【청구항 10】**

제 1 항에 있어서, 상기 비디오 스키밍을 위해 썸을 대표하는 샷 내에서 재생될 세그먼트들의 재생속도를 가변적으로 제어하는 것을 특징으로 하는 비디오 스키밍 방법.

**【청구항 11】**

제 10 항에 있어서, 단위시간당 디코딩될 프레임의 수를 정상적인 경우보다 높여서 상기 재생속도를 정상속도 보다 빠르게 재생하거나, 상기 재생할 구간 내의 프레임들을 모두 디코딩하지 않고 중간의 몇 프레임씩 스킵하여 재생함으로써 상기 재생구간을 고속 재생하는 것을 특징으로 하는 비디오 스키밍 방법.



**【청구항 12】**

제 11 항에 있어서, 상기 스킵을 이용한 고속 스키밍을 MPEG과 같은 프레임 간 압축을 사용하는 코딩 스킴을 이용한 비디오 스트림에 적용하였을 때, 디코딩할 프레임은 다른 프레임의 디코딩 없이 해당 프레임만 디코딩하여 프레임 데이터를 얻을 수 있는 I 프레임인 것을 특징으로 하는 비디오 스키밍 방법.

**【청구항 13】**

멀티미디어 데이터로서 디지털 비디오 데이터의 검색과 브라우징이 이루어지기 위하여 비디오 스키밍을 위한 사용자 명령을 입력하는 사용자 인터페이스 수단과, 상기 사용자 인터페이스수단으로 입력된 사용자 명령에 따라 해당 비디오 화일을 비디오 콘텐츠에 대한 구조적 정보로서 씬 정보를 기반으로 하고 해당 씬에 대한 샷 정보로부터 해당 씬을 대표하여 재생할 샷과 그 샷내에서 재생할 세그먼트를 선정하여 스키밍 처리하는 제어수단과, 상기 제어수단에 디지털 비디오 데이터와 해당 비디오에 대한 인덱스 정보로서 비디오 콘텐츠에 대한 씬정보 및 샷정보를 제공하기 위한 비디오 정보 화일과, 상기 제어수단에 의해서 스키밍 처리되어 논리적 단위 구조인 씬을 기반으로 하는 비디오가 재생되는 디스플레이 수단을 포함하여 이루어진 것을 특징으로 하는 비디오 스키밍 장치.

**【청구항 14】**

제 13 항에 있어서, 상기 사용자 인터페이스 수단은 비디오 스키밍의 정도로서 요약 수준을 지정하는 수단이나, 비디오 스키밍시의 재생 구간의 배속을 지


정하는 수단을 포함하여, 비디오 스키밍이 이루어질 때 비디오의 요약수준이나 재생 배속을 선택할 수 있는 것을 특징으로 하는 비디오 스키밍 장치.

【청구항 15】

제 13 항에 있어서, 상기 제어수단은 사용자의 입력 또는 기본 설정을 이용하여 스키밍 조건에 따라 인덱스 화일로부터 샷세그멘테이션 정보 및 샷 클러스터링 정보와 관련된 비디오 인덱스 정보를 읽고, 비디오 스키밍 조건에 맞게 재생할 세그먼트들을 계산하여 관련된 미디어 화일에서 해당 세그먼트들을 연속적으로 재생하여 디스플레이 수단으로 출력하는 것을 특징으로하는 비디오 스키밍 장치.

【청구항 16】

디지털 비디오 데이터, 비디오 콘텐츠의 논리적 이야기 단위 구조인 씬 정보, 비디오 콘텐츠의 물리적 편집단위인 샷 정보를 저장하는 저장수단과, 비디오 스키밍을 위하여 상기 비디오 데이터에 해당하는 상기 씬 정보를 기반으로 특정 씬을 대표하는 샷 정보를 검출하는 수단과, 상기 검출된 샷 내에서 재생할 세그먼트와 스킵할 세그먼트를 선정하는 수단과, 상기 선정된 재생할 세그먼트를 상기 저장수단에서 연속적으로 읽어서 재생하는 수단을 포함하여 이루어지는 것을 특징으로 하는 비디오 스키밍 장치.

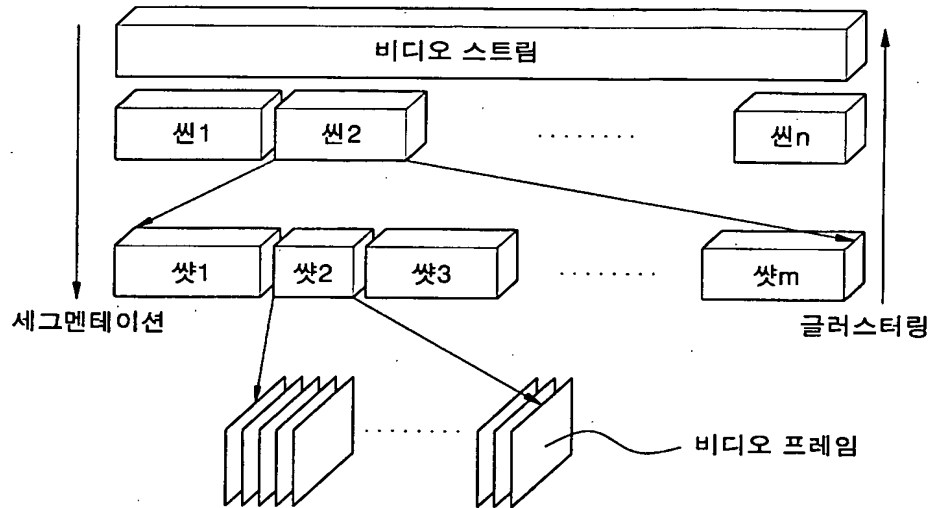


1020000055886

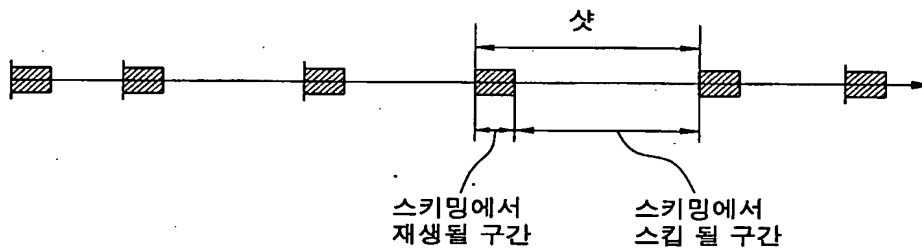
출력 일자: 2001/10/24

【도면】

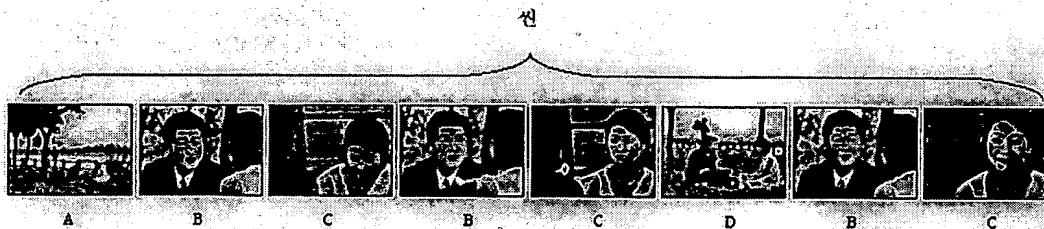
【도 1】



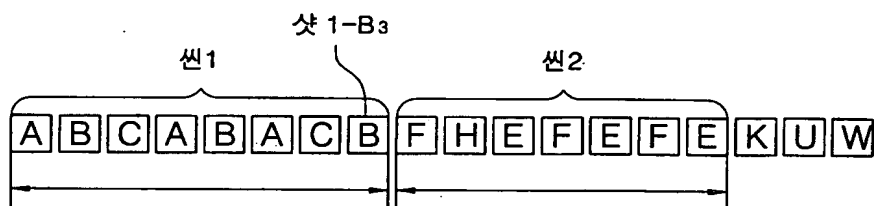
【도 2】



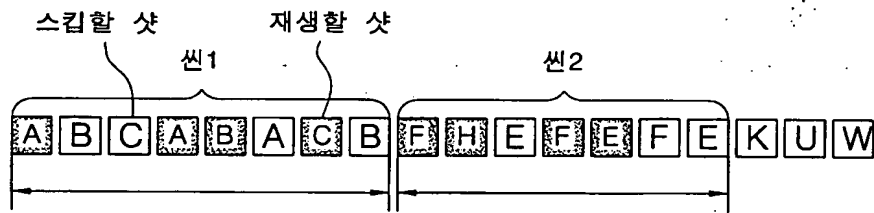
【도 3】



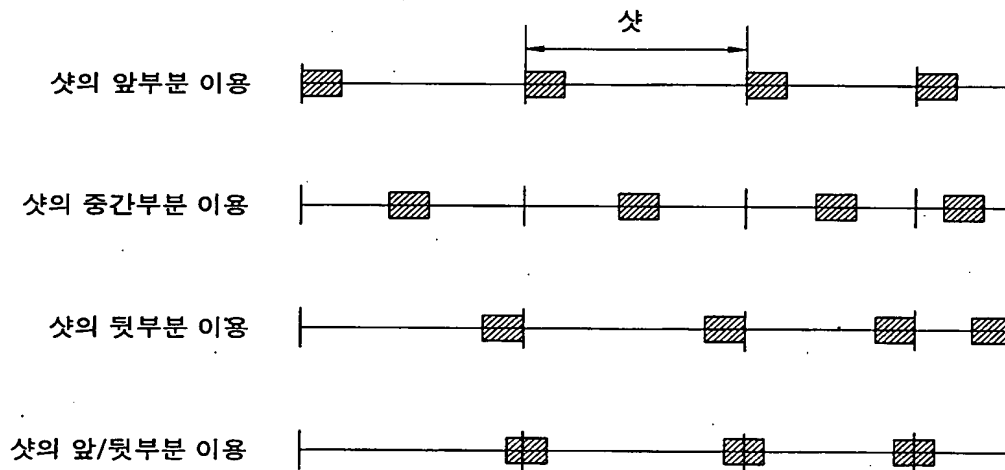
【도 4】



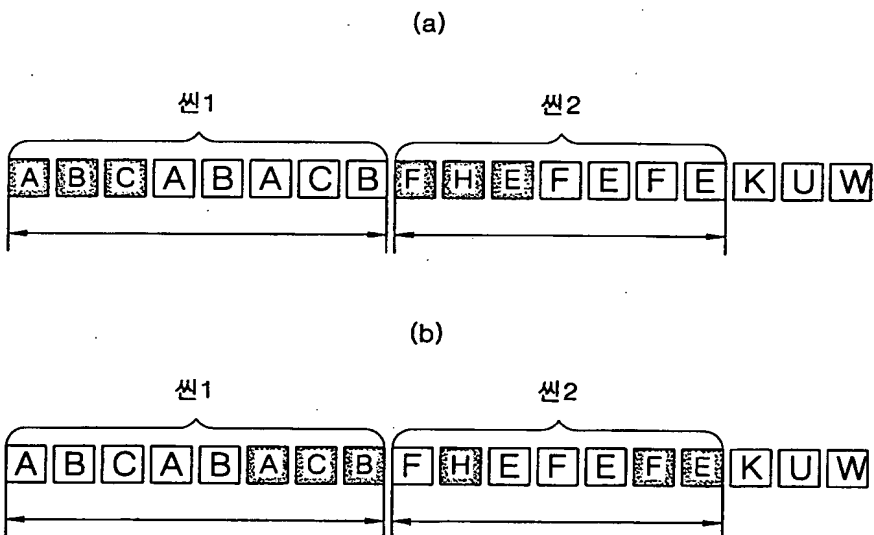
【도 5】



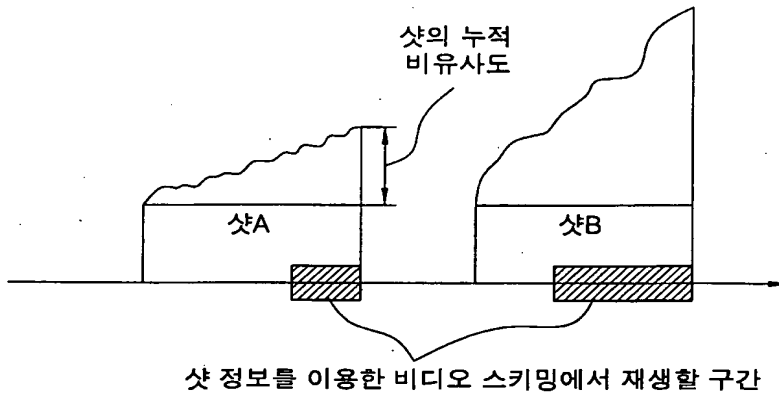
【도 6】



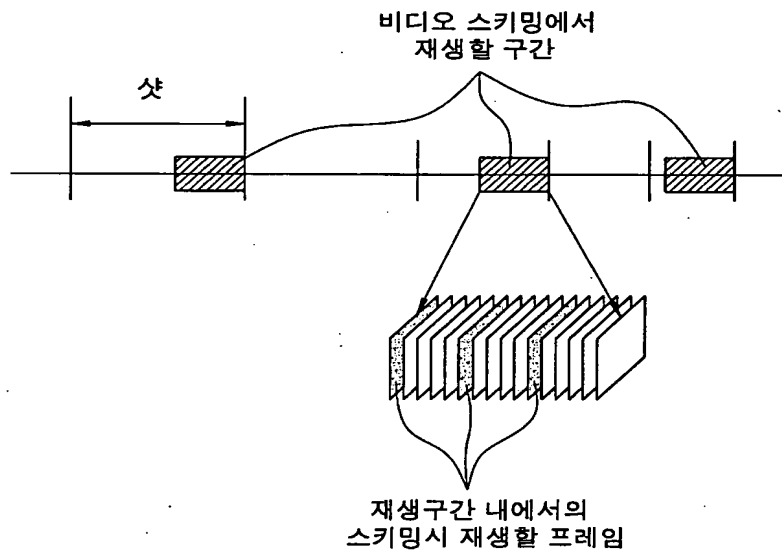
【도 7】



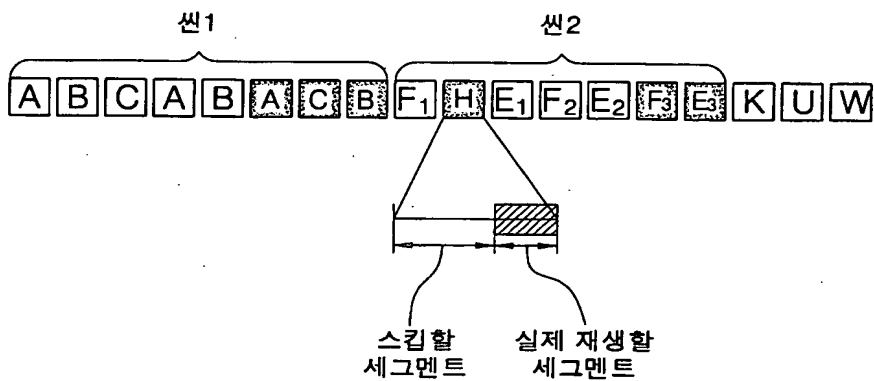
【도 8】



【도 9】



【도 10】



【도 11】

